Optimal Placement in Limit Order Book

Xin Guo
UC Berkeley

October 25, 2013, Stevens Inst. of Tech.
Outline

1 Background
   - Algorithmic trading and Limit-order-book (LOB)
   - Optimal placement vs optimal execution

2 Optimal Placement
   - The problem formulation
   - A simple correlated random walk model
   - Adding price impact and connecting with optimal execution problem
   - Some key quantities: a queuing perspective

3 Discussions
Algorithmic trading

- In an electronic order-driven market, orders arrive to the exchange and wait in the LOB (Limit Order Book) to be executed.

- An automatic and rapid trading of large quantities, with orders specified and implemented by computer algorithm.

- In US, high-frequency trading firms represent 2% of the approximately 20,000 firms operating today, but account for over 70% of all equity orders volume.

- Order flow is heavy: thousands of orders in seconds and tens of thousand of price changes in a day.
Optimal placement vs optimal execution: two time scales

- Two fundamental problems on different time scales in algorithmic tradings
- Optimal execution is for large size with daily or weekly time frame
- Optimal placement is for small size with much shorter time frame ($\approx 1/5$ min)
  Kirilenko et. al. (2011)
Optimal execution

- On a daily (weekly) time scale
- Selling a large number of shares
- Find the optimal rate of trading
- Major modeling factor: price impact
- Goal: minimize price impact or maximize expected utility

Bertimas and Lo (98), Almgren-Chriss (99,01,03), Huberman and Stanzl (00), Obizhaeva and Wang (05), Schied and Schoneborn (07,08), Almgren and Lorenz (07), Moallemi, Park and Van Roy (09), Schied, Schöneborn and Tehranchi (2010), Predoiu, Shaiket, and Shreve (2010), Weiss (2011), Gatheral and Schied (2011), Forsyth, Kennedy, Tse and Windcliff (2011), Alfonsi, Fruth, and Schied (2012), Guo and Zervos (2012)...
Given a number of (small) shares to buy or sell within a short time frame,

- using market orders or limit orders?
- the size of the order?
- what level of LOB?
- when to place the orders?
- what if the limit order can not be executed?
Recall: LOB

Six types of orders which may change the state of the limit order book

- Limit bid/ask orders: orders to buy or sell at a specified price, added to the queue and executed in order of arrival
- Market buy/sell orders: executed immediately at the best available price
- Cancellations: any unexecuted limit orders could be cancelled by their owners
- (Best) ask price: the lowest price in limit sell order
- (Best) bid price: the highest price in limit buy order
Trade off between market and limit orders

When using limit orders:
- No need to pay the spread
- Earn a rebate or discount for providing liquidity
- However no guarantee of execution: *execution risk*

When using market orders
- Have to pay the spread
- Have to pay some fees
- Yet immediate execution guaranteed

In essence: trade-off of paying the spread and fees vs execution/inventory risk
Optimal placement problem

- Buy $N$ orders before time $T > 0$ ($T \approx 1/5$ minutes)
- Split the $N$ orders in $(N_{0,t}, N_{1,t}, \ldots)$ where $N_{0,t}$ is the number of orders at market orders, $N_{1,t}$ the number of orders at the best bid, $N_{2,t}$ the number of orders at the best bid at time $t = 0, 1, \ldots, T$,
- If the limit orders are not executed by time $T$, have to buy the non-executed orders at the market price at time $T$,
- When one share of limit order is executed, the market gives a rebate of $r > 0$
- When the trader submits a market order, there is a fee of $f > 0$ for each share
- Given $(N_{0,t}, N_{1,t}, \ldots, N_{k,t})$, find the optimal strategy to minimize the overall total expected cost
Two models

- Discrete time correlated random walk model
- Continuous time: up and close with LOB
A simple correlated random walk model (G., de Larrard, Ruan (2013))

- Fix a time horizon $[0, T]$
- The spread between the best bid price ($B_t, t \geq 0$) and the best ask price ($A_t, t \geq 0$) is constant (say one tick)
- The (ask) price ($A_t, t \geq 0$) is a correlated random walk with increment of $+1$ or $-1$ tick
- No price impact: $N = 1$
The model

- The number of time step is equivalent to the number of price changes
- Take $A_n = \sum_{i=1}^{n} X_i$ with $X_i$ taking two values $-1, 1$ with transition probabilities

$$p_1 : 1 \rightarrow 1, q_1 : 1 \rightarrow -1, p_2 : -1, -1, q_2 : -1 \rightarrow 1.$$ 

where $p_1 + q_1 = p_2 + q_2 = 1.$
- $P_r(X_1 = 1) = \bar{p} = 1 - P_r(X = -1)$.
- $p_1 - p_2$ measures the “drift” of the correlated walk.
- Let $p_1 = p_2 = p$: if $p = \frac{1}{2}$, then the price is a time-changed random walk; if $p < \frac{1}{2}$, the price “mean-reverts”
For ease of exposition, let us focus on $p_1 = p_2 = p < \frac{1}{2}$ for the rest of the talk.

- $P(A(n) = k) = P(\sum_{i=1}^{n} X_i = k)$
- $Y_n = \min\{A_m, m \leq n\}$
- Diffusion limit

Mohan (1955), Renshaw and Henderson (1981)
Step I: Optimal strategy for one period model

- Orders can only be placed in the beginning
- Place market orders to fill the position at the end of the period
- A limit order placed at price $-k$ will be executed with probability one if $Y_T \leq -k$, and with probability $q$ if $Y_T = -k + 1$.
- Analyzing $C_i$: the average cost of place 1 share at $i$-th limit from the ask

**Proposition**

The optimal strategy for one period model is using either the market order or the best bid. The decision is a threshold type, depending on $r$, $f$, $\bar{p}$ and $p$. In particular, if $\bar{p} = \frac{1}{2}$, then it is the best to peg the bid price.
Step II: Optimal placement for multi-period

- Only focus on the best bid and ask queues
- Allowed to change any non-executed orders every time after the price moves, without cost
- After each price change, decisions for non-executed orders are: place it on the new market order, or place it on the new limit order, or cancel the unexecuted limit order and wait
- No fees for order cancelation after the best ask price changes
- An order in a limit order will be executed with probability $q$ before next price move
Optimal strategy for multi-period model

Theorem

- $0 \leq s \leq T$ be the remaining time to go
- If $0 \leq s < T$: when the previous price change is going up: there exists $s_1^* \in [2, T]$ so that then for any $s \leq s_1^*$, use limit order; if $s \geq s_1^*$, wait.
- If $0 \leq s < T$: when the previous price change is going down: there exists $s_2^* \in [1, \infty]$, so that then for any $s \leq s_2^*$, use market order; when $t \geq s_2^*$, use limit order.
- $s_1^* > s_2^* > 0$
- If $s = T$: there exist two thresholds $\bar{p}_1^* \leq \bar{p}_2^* \in [0, 1]$ such that when $\bar{p} < \bar{p}_1^*$, wait; when $\bar{p}_1^* \leq \bar{p} \leq \bar{p}_2^*$, use limit order; otherwise, use market order.
Illustration

The previous price change is up:
- The previous price change is up
- T
- S_1^*
- Best bid

The previous price change is down:
- The previous price change is down
- T
- S_2^*
- Market order
- Best bid
Remark

- The threshold-type optimal strategy is expected given the Markov property of \((A_t, A_{t-1})\) or equivalently that of \((A_t, X_t)\).
- The optimal strategy is in fact independent of \(A_t\): in the optimal placement problem, the best strategy is derived from comparing choices of the best bid, the market order or no action at any given price level \(A_s\), hence the absolute expected value of each strategy is less relevant.
Step III: adding price impact

- $N \geq 1$
- adding linear price impact for both limit order and market order
- martingale model replacing the simple correlated random walk model

**Theorem** For each period, use either limit order or market order, but not both. In the particular case of $q = 0$ or $q = 1$, the optimal strategy is the same as in Bertsimas and Lo (1998) for the optimal execution problem.
Related works

- Hult and Kiessling (2010)
- Adding the possibility of intermediate selling/market making
- Optimal placement across exchanges
  Laruelle, Lehalle, and Pagès (2011) and Cont and Kukanov (2012)
Learning from the toy model

Key quantity: the probability of an order at some level of LOB being executed by time $T$, when placed at time $t$. It depends on the queue length, its position in the LOB, frequency of price changes.
Up and close with LOB
Focus on $Z(t)$

- Let $(Z_t, t \geq 0)$ be the position of one particular order in the bid queue.
- The order is executed at time $\tau$ for $\tau := \inf\{t \geq 0, \ Z_t = 0\}$.
- $Z(t)$ contains crucial information of the balance and relation between the execution risk and the microstructure of the LOB.
- Starting point: focus on the best bid/ask. That is, the limit order placed on best bid/ask.
Queue Tail

A limit bid order

A particular position

A market sell order

Queue Head

Figure: Orders happened in the best bid queue
Model assumptions and notations

- Focus on the best bid and best ask
- Six types of orders: limit orders, market orders, and cancelations at best bid queue and best ask queue.
- All orders are point processes. For example, $N_{lb}(t)$, the arrival process of the limit orders at the best bid queue up to time $t$ and $V_{lb}'$ is the size of the $i'$th limit orders at the best bid queue.
- Arrival processes are mutually independent, and independent of order sizes
  - $\lambda_{lb}, \lambda_{la}, \lambda_{mb}, \lambda_{ma}, \lambda_{cb}, \lambda_{ca}$: the corresponding intensities of the arrival processes.
  - $\bar{V}_{lb}, \bar{V}_{la}, \bar{V}_{mb}, \bar{V}_{ma}, \bar{V}_{cb}, \bar{V}_{ca}$: the means of the corresponding order sizes of different types.
  - $v^2_{lb}, v^2_{la}, v^2_{mb}, v^2_{ma}, v^2_{cb}, v^2_{ca}$: the variances of the corresponding order sizes of different types.
Diffusion limit

Theorem

Let \( Q^b_n(t) \), \( Q^a_n(t) \), and \( Z_n(t) \) be appropriately rescaled queue sizes and the order position. Assuming the cancellation size is proportional to the queue length. Then under proper mixing conditions,

\[
(Q^b_n(t), Q^a_n(t), Z_n(t)) \xrightarrow{n} (Q^b(t), Q^a(t), Z(t))
\]

where \( Q^b(t) \), \( Q^b(t) \) are “regulated” planar Brownian motion, and \( Z(t) \) takes the form of

\[
dZ(t) = -\left(\lambda_{mb} \bar{V}_{mb} + \lambda_{cb} \bar{V}_{cb} \frac{Z(t)}{Q^b(t)}\right)dt + \sqrt{\lambda_{mb}(v_{mb})^2 + \lambda_{cb}(v_{cb})^2 \frac{Z(t)}{Q^b(t)}}dW(t)
\]

with integral-differential boundary conditions.
Step 1: proper time rescaling

\[ Q^b_n(t), Q^a_n(t), Z_n(t): \text{the scaled queues characterized by the sequences of} \]
\[ LB_n(t), MB_n(t), CB_n(t), LA_n(t), MA_n(t), CA_n(t). \text{For instance,} \]
\[ LB_n(t) = \sum_{i=1}^{N_{lb}(nt)} \left[ \frac{V_{lb}^i - \bar{V}_{lb}}{\sqrt{n}} + \frac{\bar{V}_{lb}}{n} \right] \]
Step 2: diffusion limit similar to R. Cont and A. de Larrard (2010)

\[(Q_n^b(t), Q_n^a(t)) \xrightarrow{n} (Q^b(t), Q^a(t))\]

where \(Q^b(t), Q^b(t)\) constitute a regulated planar Brownian motion
Quantities

$\delta$

$s_b$

$s_a$

$Q_b$

$Q_a$
Step 3: Convergence theorem for the SDE’s

Kurtz and Protter (1991)

Let $X_n(t) = U_n(t) + \int_0^t F_n(X_n(s), s) dY_n(s)$, where $U_n$ and $Y_n$ are “nice”.

Suppose $(U_n, Y_n) \to (U, Y)$, $(x_n, y_n) \to (x, y)$ implies $(x_n, y_n, F_n(x_n)) \to (x, y, F(x))$. Then under some technical conditions any limit point of the sequence $X_n$ satisfy the SDE

$$X(t) = U(t) + \int_0^t F(X, s) dY(s)$$
In our case, let

\[ Y_n = CB_n \]
\[ X_n = (Q_n^b, Q_n^a, Z_n) \]
\[ F_n = (0, 0, \frac{Z_n}{Q_n^a}) \]
\[ U_n = (LB_n - MB_n - CB_n, LA_n - MA_n - CA_n, -MB_n) \]
Related to queuing theory

(Harrison, Rieman, Williams, Whitt ...)

- $Z(t)$ vs. the “workload process”: classical queuing concerns on status/stability of the system, while algorithm trading focuses on the individual trade


- Cancelation rate in LOB is high and motivation and characteristics of cancelations are unclear. In fact, different assumptions of cancelation lead to different forms of diffusion approximation for $Z(t)$
Most relevant statistical issues for optimal placement

- Understanding and identifying characteristics/motivation of cancelation in LOB
- Testing “mean-reverting” in LOB
- Identifying order book imbalance
- Identifying hidden liquidity (dark pool and non-displayable orders)
- Understanding different level of liquidity in different market: diffusive or not?
- Large data analysis with Lawrence Berkeley National Lab and NASDAQ OMX Group
This talk is based on

THANK YOU